



48649/97

PCT

VERTRAG ÜBER DIE  
TENTWESENS (PCT)

(51) Internationale Patentklassifikation 6 : <b>C12N 15/30, C07K 14/445, C12N 15/62, A61K 39/015, 31/70, C07H 21/00, C12N 1/21, 5/10, 1/11, 15/67</b>		A3	(11) Internationale Veröffentlichungsnummer: <b>WO 98/14583</b>
		(43) Internationales Veröffentlichungsdatum:	9. April 1998 (09.04.98)
(21) Internationales Aktenzeichen: <b>PCT/EP97/05441</b>		(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) Internationales Anmeldedatum: <b>2. Oktober 1997 (02.10.97)</b>		Veröffentlicht <i>Mit internationalem Recherchenbericht. Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist. Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i>	
(30) Prioritätsdaten: <b>196 40 817.2                      2. Oktober 1996 (02.10.96)                      DE</b>		(88) Veröffentlichungsdatum des internationalen Recherchenberichts: <b>16. Juli 1998 (16.07.98)</b>	
(71)(72) Anmelder und Erfinder: <b>BUJARD, Hermann [DE/DE]; Remlerstrasse 9, D-69120 Heidelberg (DE).</b>			
(72) Erfinder; und			
(75) Erfinder/Anmelder (nur für US): <b>TOLLE, Ralf [DE/DE]; Friedrich-Naumann-Strasse 8, D-71636 Ludwigsburg (DE). PAN, Weiqing [CN/DE]; Im Buschgewann 71, D-69123 Heidelberg (DE).</b>			
(74) Anwalt: <b>GRÜNECKER, KINKELDEY, STOCKMAIR &amp; SCHWANHÄUSSER; Maximilianstrasse 58, D-80538 München (DE).</b>			
(54) Title: <b>METHOD FOR PRODUCING RECOMBINANTS INTENDED FOR USE IN A COMPLETE MALARIA ANTIGENE GPI90/MSP1</b>			
(54) Bezeichnung: <b>REKOMBINANTES HERSTELLUNGSVERFAHREN FÜR EIN VOLLSTÄNDIGES MALARIA-ANTIGEN GPI90/MSP1</b>			
(57) Abstract			
<p>The present invention relates to a method for producing recombinants intended for use in the complete cell-surface protein gp190/MSP1 from plasmodium, especially plasmodium falciparum, as well as the complete DNA sequence of this protein and the appropriate host organisms suited for expressing said sequence, whereby the protein concerned can be entirely synthesized outside the parasite. Also, the inventive method enables sufficient production of above-mentioned protein and its supply as a vaccine. Finally disclosed is a process for stabilizing genes with high At concentration.</p>			
(57) Zusammenfassung			
<p>Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das komplette gp190/MSP1-Oberflächenprotein von Plasmodium, insbesondere Plasmodium falciparum, sowie die vollständige DNA-Sequenz dieses Proteins und geeignete Wirtsorganismen für die Expression der Sequenz, wodurch das Protein in seiner Gesamtheit erstmals außerhalb des Parasiten synthetisiert werden konnte. Die Erfindung eröffnet erstmals die Möglichkeit, das gp190/MSP1-Oberflächenprotein in ausreichender Menge herzustellen; ferner ist es ein Gegenstand der Erfindung gp190/MSP1 als Impfstoff zur Verfügung zu stellen. Schließlich gibt die Erfindung ein Verfahren zur Stabilisierung von AT-reichen Genen an.</p>			

\*(Siehe PCT Gazette Nr. 32/1998, "Section II")

## **Recombinant process for preparing a complete malaria antigen, gp190/MSP1**

The invention concerns a recombinant manufacturing process for the complete malaria antigen gp190/MSP1, as well as separate naturally-occurring domains and parts of the same, by expression of a synthetic DNA sequence. The invention concerns in addition the DNA sequences produced by the process and the host organisms used for the expression of the DNA sequences. In addition the invention concerns the use of the complete malaria antigen as well as parts thereof as a vaccine for immunization against malaria.

Finally the invention under consideration concerns a stabilization process for AT-rich genes, as well as stabilized genes which are characterized by a reduced AT content.

Malaria is one of the most significant infectious diseases in the world. According to WHO reports, in 1990 40% of the world population in 99 countries was exposed to the risk of malaria. At the same time its distribution is enormously on the increase. This may be ascribed above all to intensive development of resistance in the parasites causing malaria, promoted by the recommendation and use as prophylactics of the drugs intended for treatment. Besides the search for new and effective chemotherapeutic agents hope is nowadays directed towards the development of vaccines, since people in areas of the world where malaria is epidemic do manage to develop some kinds of immunity. As well as a natural resistance to malaria, such as that found in heterozygous carriers of the sickle-cell gene and people with thalassaemia and glucose-6-phosphate dehydrogenase deficiency, in the course of malarial infection in humans immune mechanisms can be stimulated which express themselves in a heightened capacity for resistance to the Plasmodia. Consequently the course of the disease in populations exposed to severe epidemics is generally less threatening than in persons exposed to the infection less frequently or for the first time.

The main problem in the development of a vaccine is the identification of an antigen which can induce protective immunity, since there is no easily accessible well-defined animal model available for the four parasites affecting man. The organism causing malaria belongs to the Plasmodium group, of which infection with one of the four parasites Plasmodium vivax,



Plasmodium ovale, Plasmodium malariae or Plasmodium falciparum results from the bite of Anopheles mosquitoes. Of these parasites Plasmodium falciparum is the most dangerous and the most widely distributed.

The main surface protein of the merozoite, the invasive form of the blood stage of the malaria parasite Plasmodium falciparum and other malaria parasites such as P. vivax, is a 190-220 kD glycoprotein. Late in the development of the parasite this precursor is processed into smaller proteins, which can however be isolated from merozoites as a unitary complex. By means of a glycosylphosphatidyl-inositol bond this complex is coupled to the merozoite membrane. The sequences of the gp190 proteins of various P. falciparum strains fall into two groups, between which intragenic recombination is frequent. In general the protein consists of many highly conserved regions, of a dimorphic zone to which in each case one of two alleles belongs, and of two relatively small oligomorphous blocs in the N-terminal zone (Tanabe, K., Mackay, M., Goman, M. and Scaife, J.G. (1987), Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmodium falciparum. J. Mol. Biol. 195, 273-287, Miller, L.H., Roberts, T., Shahabuddin, M. and McCutchan, T.F. (1993), Analysis of sequence diversity in the Plasmodium falciparum merozoite surface protein-1 (MSP-1). Mol. Biochem. Parasitol. 59, 1-14).

Already early on gp190/MSP1 was considered as a possible candidate for a vaccine. In the rodent model active protection against infection with rodent parasites was obtained following immunization with the analogous protein. Passive protection could be procured with antibodies directed against this protein (see also Holder, A.A. and Freeman, R.R. (1981), Immunization against blood-stage rodent malaria using purified parasite antigens, Nature 294, 361-364; Marjarian, W.R., Daly, T.M., Weidanz, W.P. and Long, C.A. (1984), Passive immunization against murine malaria with an IgG3 monoclonal antibody, J. Immunol. 132, 3131-3137). The data which ought to support this assumption are nevertheless in details not statistically significant.

There are in addition, a number of monoclonal antibodies which in vitro inhibit the invasion of erythrocytes by P. falciparum and are directed against gp190/MSP1 (Pirson, P.J. and Perkins, M.E. (1985), Characterization with monoclonal antibodies of a surface antigen of Plasmodium



*falciparum* merozoites. J. Immunol. 134, 1946-1951; Blackman, M.J., Heidrich, H.-G., Donachie, S., McBride, J.S. and Holder, A.A. (1990), A single fragment of a malaria merozoite surface protein remains on the parasite during red-cell invasion and is the target of invasion-inhibiting antibodies. J. Exp. Med. 172, 379-382).

Finally, a series of vaccine studies have been carried out with gp190/MSP1 from *P. falciparum* on primates, particularly on Aotus and Saimiri monkeys (see also Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. and Richle, R. (1984), Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual blood stages, J. Exp. Med. 160, 441-451; Hall, R., Hyde, J.E., Goman, M., Simmons, D.L., Hope, I.A., Mackay, M. and Scaife, J.G. (1984), Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria, Nature 311, 379-382; Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. and Kan, S.-C. (1987), Merozoite surface coat precursor protein completely protects Aotus monkeys against Plasmodium falciparum malaria, Proc. Natl. Acad. Sci. USA 84, 3014-3018; Ettlinger, H.M., Caspers, P., Materile, H., Schoenfeld H.-J., Stueber, D. and Takacs, B. (1991), Ability of recombinant or native proteins to protect monkeys against heterologous challenge with Plasmodium falciparum, Inf. Imm. 59, 3498-3503; Holder, A.A., Freeman, R.R. and Nicholls, S.C. (1988), Immunization against Plasmodium falciparum with recombinant polypeptides produced in Escherichia coli, Parasite Immunol. 10, 607-617; Herrera, S., Herrera, M.A., Perlaza, B.L., Burki, Y., Caspers, P., Doebeli, H., Rotmann D. and Certa, U. (1990), Immunization of Aotus monkeys with Plasmodium falciparum blood-stage recombinant proteins, Proc. Natl. Acad. Sci. USA 87, 4017-4021; Herrera, M.A., Rosero, F., Herrera, S., Caspers, P., Rotmann, D., Sinigaglia, F. and Certa, U. (1992), Protection against malaria in Aotus monkeys immunized with a recombinant blood-stage antigen fused to a universal T-cell epitope; correlation of serum gamma interferon levels with protection, Inf. Imm. 60, 154-158; Patarroyo, M.E., Romero, P., Torres, M.L., Clavijo, P., Moreno, A., Martinez A., Rodriguez, R., Guzman, F. and Cabezas, E. (1987), Induction of protective immunity against experimental infection with malaria using synthetic peptides, Nature 328, 629-632). In these vaccine studies two premises may be distinguished:

- Use of material isolated from parasites, and



- Administration of material procured in heterologous systems of expression.

The latter consists as a rule of relatively small segments of the total protein. Although the results of the inoculations carried out preliminarily on monkeys indicate that gp190/MSP1 could bring about protection, all the experiments carried out on primates have two problems, which place such a conclusion in question:

- (a) they were carried out on too small groups of animals
- (b) they were not repeated.

The results and the conclusions drawn from them are consequently not statistically confirmed. Besides the difficulty of access to suitable monkeys there remains the main basic problem, that it has so far not been possible to manufacture good vaccination material in a suitable quantity.

On the other hand, after the sequencing of the gp190 gene from the K1 and MAD20 strains of *Plasmodium falciparum* overlapping fragments could be expressed in *E. coli*. With this material epidemiological studies in West Africa showed that in the adolescent group a correlation existed between antibody titre against gp190/MSP1 fragments on one hand and protection from parasite infection on the other. In addition the titre also appeared to correlate with the capacity to control the parasitaemia even at a low level (Tolle et al. (1993): A prospective study of the association between the human humoral immune response to *Plasmodium falciparum* blood stage antigen gp190 and control of malarial infections, *Infect. Immun.* 61, 40-47). These results are supplemented by new investigations on Aotus monkeys in the framework of the present invention. Here an enhanced protection against infection with the parasite was attained because protein preparations from *Plasmodium falciparum*, which consisted predominantly of unprocessed gp190/MSP1, had been used as vaccine. The monkeys with the highest antibody titres against gp190/MSP1 were the best protected. These results eventually indicated gp190 as a most promising candidate for a vaccine against tropical malaria.



By some groups of workers the C-terminal domain of gp190 (p19 or p42) is assigned a particular role in the immunity mediated by gp190 (see also Chang, S.P., Case, S.e., Gosnell, W.L., Hashimoto, A., Kramer, K.J., Tam, L.Q., Hashiro, C.Q., Nikaido, C.M., Gibson, H.L., Lee-Ng, C.T., Barr, P.J., Yokota, B.T. and Hui, G.S.N. (1996), A recombinant baculovirus 42-kilodalton C-terminal fragment of Plasmodium falciparum merozoite surface protein 1 protects Aotus monkeys against malaria, Inf. Imm. 64, 253-261; Burghaus, P.A., Wellde, B.T., Hall, T., Richards, R.L., Egan, A.F., Riley, E.M., Ripley-Ballou, W. and Holder A.A. (1996), Immunization of Aotus nancymai with recombinant C-terminus of Plasmodium falciparum merozoite surface protein 1 in liposomes and alum adjuvant does not induce protection against a challenge infection, Inf. Imm., in press.

Thus far, however, it has also been impossible to exclude other parts of gp190 on a rational basis as irrelevant to a protective immune response. Hence it is as necessary as ever to use the entire gene or the intact gp190 for vaccine investigations. Despite multiple investigations by various work-groups, however, there has not yet been any success in cloning and expressing the entire gp190/MSP1 gene.

Nor has it so far been possible to exclude a priori any part of the gp190 sequence as irrelevant to the protective immune response, so that it is as necessary as ever to use the entire gene or gene product for vaccine investigations. Nevertheless, despite many investigations by a number of working groups there has not yet been any successful cloning of the whole gene for gp190/MSP1.

The present invention provides a means of making available an adequate quantity of vaccine material in the form of the complete gp190/MSP1. The present invention



further provides a process by which this vaccine material could be recovered.

In addition the present invention provides a  
5 complete DNA sequence of gp190/MSP1 which could be expressed in a host organism.

The present invention also provides a host  
10 organism containing the complete gp190/MSP1 gene.

Finally, the present invention provides a  
stabilisation process for AT-rich genes, as well as a  
stabilised gene suitable for expression characterised in a  
reduction of the AT content.  
15

In the following, certain concepts are explained  
in more detail in order to make clear how they should be  
understood in this context.

20 Throughout the description and claims of this specification, the word "comprise" and variations of the word, such as "comprising" and "comprises", means "including but not limited to", and is not intended to exclude other additives, components, integers or steps".  
25

"Recombinant manufacturing process" means that a protein of a DNA sequence is expressed by a suitable host organism in which the DNA sequence has arisen from cloning and fusion of individual DNA fragments.  
30

"Complete gp190/MSP1 protein" here means the entire gp190/MSP1 surface protein isolatable from the above named Plasmodia, especially Plasmodium falciparum, representing the main surface protein of the above named parasite as well as the proteins with analogous function from the Plasmodium species such as P. vivax. The term  
35 therefore comprises in each case the main surface protein



of the merozoites of the four malaria parasites named above as dangerous to man. "Complete gp190/MSP1 gene" means the gene coding for this protein. In this context "complete" signifies that the entire amino-acid sequence of the native protein is present or that the gene sequence codes for the entire amino-acid sequence of the native protein. Mutated and/or shortened forms of gp190/MSP1 are however included therewith insofar as they display the same immunisation potential (vaccine protection) as the complete gp190/MSP1. Finally the term also includes variants of gp190/MSP1 characterised by containing in one molecule protein fragments of various alleles.

"FCB-1" is a strain of *P. falciparum* such as that described in Heidrich, H.-G., Miettinen-Baumann, A., Eckerskorn, C. and Lottspeich, F. (1989) The N-terminal amino acid sequences of the *Plasmodium falciparum* (FCB1) merozoite surface antigens of 42 and 36 kilodalton, both derived from the 185-195-kilodalton precursor. Mol. Biochem. Parasitol. 34, 147-154.

"Attachment signal" here means a protein structure for by a DNA sequence at the 3' or 5' end of the gene according to the invention. Attachment signals are structures enabling the attachment of a polypeptide to other structures, such as for example membranes.

"Signal peptide" here signifies a protein structure coded for which a DNA sequence at the N-terminal end of the gene according to the invention codes. Signal peptides are structures which among other things enable penetration of the polypeptide into membranes.

In the context of the present invention "AT-content" means the percentage amount of adenine-thymine base pairs compared to guanine-cytosine base pairs.



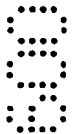
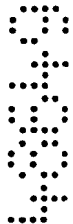


"Cloning" will comprehend here all known state-of-the-art cloning methods which could be applied here, but which are nevertheless not all described in detail because  
5 they belong among the normal tools of the person skilled in the art.

"Expression in an appropriate expression system" should here include all known state-of-the-art methods of  
10 expression in known expression systems which could be applied here, but which are nevertheless not all described in detail because they belong among the normal tools of the person skilled in the art.

15 The present invention provides a process by which the protein gp190/MSP1 and its gene can be produced in sufficient quantity without excessive cost.

For the first time it is possible by this process  
20 to synthesise the protein in its entirety outside the parasite. As the analysis with conformational epitope-recognising monoclonal antibodies shows, the protein thus synthesised is at least reproducibly synthesisable over wide areas in naturally folded form. By the recombinant  
25 manufacturing process many milligrams of intact gp190/MSP1 could in every case be recovered from the host organism, a quantity which for



technical and economic reasons can never be recovered from parasites. Production of the protein in any desired quantity is now possible and opens new perspectives for its use as an experimental vaccine against malaria. Furthermore, the way is now open for the development of living vaccines as well as for vaccines based on nucleic acids.

Synthesis of the gene sequence coding for the protein gp190/MSP1 is preferentially based on the sequence of the FCB-1 strain of *P. falciparum*. *P. falciparum* is the agent of tropical malaria and hence of the most dangerous among the types of malaria. The basic gene is a representative of the "K1 allele", where K1 stands for a particular *P. falciparum* strain. Its coding sequence extends over 4917 base pairs and includes a signal sequence at the N-terminal end as well as an attachment sequence at the C-terminal end.

Furthermore, according to the invention the recombinant manufacturing process is preferentially characterized in having the AT content of the DNA sequence on which the protein is based reduced relative to the wild type, from 74% in the original gene preferably to about 55%, for example while the amino-acid sequence of the FCB-1 protein is maintained a DNA sequence with the codon frequencies usual in the human genome is produced. Other codon frequencies which reduce the AT content are also conceivable.

Preferentially the gene underlying the protein produced by the recombinant manufacturing process codes for the full amino-acid sequence including signal peptide and GPI attachment signal peptide, further described as gp190<sup>s</sup>.

In another preferred embodiment, the gene on which the protein produced by the recombinant manufacturing process is based codes for the complete amino-acid sequence except for the GPI attachment signal. This embodiment is then described as gp190<sup>s1</sup>.

In yet another preferred embodiment, the gene on which the protein produced by the recombinant manufacturing process is based codes for the complete amino-acid sequence except for the GPI attachment signal and the signal peptide. This embodiment is then described as gp190<sup>s2</sup>.



In a further preferred embodiment type, the gene on which the protein produced by the recombinant manufacturing process is based codes for the complete amino-acid sequence and a trans-membrane attachment sequence.

In a particularly preferred embodiment the recombinant manufacturing process includes the following steps:

In the first place the design of the DNA sequence to be synthesized on the basis of the gene from *P. falciparum* FCB-1, in which a DNA sequence with for example the codon frequencies common in the human genome is manufactured with retention of the amino-acid sequence of the FCB-1 protein.

The AT content of the gene should be reduced by this, preferably to 55%. Further on in the process the planned sequence is divided for example into 5 overlapping regions, which at the same time correspond to domains of the natural processing products of gp190/MSP1 from FCB-1: p83, p31, p36, p30 and p19.

Desoxyoligonucleotides are synthesized, which in each case extend the entire length of a region.

The desoxyoligonucleotides so synthesized are particularly preferred where their sequence corresponds in an alternating manner to the "upper" (5' - 3') or the "lower" (3' - 5') DNA strand. The length of these oligonucleotides is preferably on average 120 nucleotides and they overlap the neighboring sequences in each case by about 20 bases.

In one possible embodiment DNA sequences of about double the length of the existing end-products are manufactured by asymmetrical PCR, in effect so that the superfluous DNA sequences nearby in each case represent the opposite strand. This leads in a second

PCR amplification cycle to a second product corresponding to the length of four originally inserted oligonucleotides excluding the overlapping region. Transfer of these products to a preparation consisting predominantly of single-stranded DNA by asymmetrical PCR with the



terminal oligonucleotides permits the manufacture in a further amplification step of an 800-bp long double-stranded DNA fragment in only 25 PCR cycles.

5 In this manner the regions coding for p19, p30, p36 and p31 are directly synthesised and molecularly cloned in *E. coli*. Clones with fault-free sequences are conserved either directly or by the joining up of fault-free sequence fragments. The region which codes for p83 is constructed  
10 by fusion from two sequences comprising about 1200 bp.

In the further course of production single sequences are cloned. As expression vectors candidates preferred are the plasmids pDS56, RBS11 ("Hochuli, E.,  
15 Bannwarth, W., Doebeli, H., Gentz, R. and Stueber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Biotechn.* 6, 1321-1325"), pBi-5 ("Baron, U., Freundlib, S., Gossen, M. and Bujard, H. (1995) Corregulation of two gene  
20 activities by tetracycline via a bidirectional promoter. *Nucl. Acids Res.* 23, 3605-3606") and ppTMCS. It is possible nonetheless also to conceive of other expression vectors.

25 Host organisms preferred for expression are *E. coli*, with the strain DH5alphaZ1 specially preferred (R. Rutz, Dissertation 1996, Heidelberg University), HeLa cells, CHO cells, *Toxoplasma gondii* (Pfefferkorn, E.R. and Pfefferkorn, C.C. 1976, *Toxoplasma gondii*: Isolation and  
30 preliminary characterisation of temperature-sensitive mutants. *Exp. Parasitol.* 39, 365-376) or *Leishmania*. Additional host systems might be e.g. yeasts, baculoviruses or adenoviruses, so that the subject matter of the invention should not be limited to the host systems  
35 mentioned.

The present invention also provides a complete



DNA sequence, suitable for expression, of the gp190/MSP1 surface protein of *P. falciparum*.

In a preferred embodiment of the present  
5 invention the sequence suitable for expression codes for the complete amino-acid sequence.

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
22

In another preferred embodiment of the present invention the sequence suitable for expression codes for the complete amino-acid sequence except for the attachment signal.

In a further preferred embodiment according to the present invention the DNA sequence suitable for expression codes for the complete amino-acid sequence except the attachment signal and the peptide signal. This embodiment of gp190/MSP1 can hence be characterized in including at the N-terminus 11 additional amino-acids, of which 6 are histidines.

Particularly preferred the DNA sequence suitable for expression contains no recognizable "splice-donor" and "splice-acceptor" sites, and is preferably characterized in not containing any larger GC-rich sequences which might result in stable hairpin structures at the RNA level.

Recognition signals for restriction enzymes which recognize sequences of six or more base pairs should preferably be avoided.

In a preferred embodiment specific cleavage sites for restriction endonucleases, occurring only once in the gene, are introduced into regions to separate the existing domains following processing of the protein.

Particularly preferred would be the presence at both ends of the gene of sequences for restriction endonucleases which do not occur in the gene.

Furthermore host organisms containing the complete sequence of gp190/MSP1 surface protein are provided by the invention.

Such host organisms are preferably E. coli, particularly preferred being the strain DH5alphaZ1, HeLa cells, CHO cells, Toxoplasma gondii or Leishmania. The HeLa and CHO cells ought preferably to synthesize constitutively tTA.



Finally the present invention provides a possibility of using a gp190/MSP1 surface protein created produced according to the recombinant manufacturing process, or parts thereof, for active immunization against malaria.

The scheme for synthesis presented here also permits manufacture of the second allele of the gp190/MSP1 gene, whereby the dimorphism of the protein is also taken into account. The main variability of the protein depends however on the sequences of two relatively short blocs, blocks II and IV (ref. 1), which are oligomorphous. The present sequence data make it possible to disclose over 95% of all known gp190/MSP1 sequences with 6-8 sequence combinations of these blocs. The synthesis of these sequence variants can be brought about problem-free by means of the strategies proposed here, so that variants can be built up both in the K1 and in the MAD20 allele. Vaccines from the families of sequences thus created can confer protection where required against a wide spectrum of parasites with gp190/MSP1 variants.

The manufacture of different types of vaccine is possible:

- At the level of protein preparations, where in each instance mixtures of the two families (K1 type, MAD20 type with different variants of Blocs II and IV) can come into use. Various carrier or adjuvant materials could be added: aluminum oxide, liposomes, IscomsQSz1, etc.

- At the level of live vaccines: (a) viral carriers, especially vaccinia and adenoviruses; (b) parasites as carriers, particularly avirulent forms of Leishmania and Toxoplasma; (c) bacterial carriers, e.g. Salmonella.

- At the level of nucleic acids, whereby for example vectors suitable for gene therapy would be used to introduce the gene into the host; beyond that the introduction of nucleic acids coding for the desired protein can be envisaged.

A further possibility for vaccination lies in the use of a gp190/MSP1 protein produced according to the recombinant manufacturing process set out by the invention, for the production of monoclonal antibodies which can then be used in their turn for passive immunization against malaria.



Similarly it becomes possible to use the DNA sequence on which the protein is based at an intermediate stage arising in the course of the recombinant manufacturing process for the construction of a vaccine based on nucleic acids.

Finally the invention also concerns a process for the stabilization of gene sequences, especially for sequences which do not show adequate stability in expression systems.

According to the invention this stabilization is attained because the AT content of the sequence is reduced.

Moreover a stabilized gene characterized by having a reduced AT content is provided by the invention. An example of such a stabilized gene is the gene for gp190/MSP1 surface protein according to the present invention.

In the following the invention will be described with the help of figures and tables as well as some examples in individual embodiments.

They show:

Fig. 1: Schematic representation of the gp190/MSP1 precursor protein from *P. falciparum* (FCB-1).

Fig. 2: Two vaccine trials carried out on Aotus monkeys with native gp190/MSP1 from *P. falciparum* (FCB-1).

Fig. 2A: With 3 x 60 micrograms gp190/MSP1

Fig. 2B: With 3 x 40 micrograms gp190/MSP1

Fig. 3A: Strategy of synthesis of the gp190/MSP1 gene





Fig. 3B: Principle of PCR-based total synthesis

Fig. 3C: Total sequencing of gp190<sup>S</sup>

Fig. 3D: N- and C-termini of gp190<sup>S1</sup> variant

Fig. 4A: Expression vector pDS56 with gp190<sup>S2</sup> sequence

Fig. 4B: Gel electrophoresis of gp190<sup>S2</sup>

Fig. 5A: Expression vector pBi-5 with gp190<sup>S1</sup> sequence

Fig. 5B: Immunofluorescence of HeLa cells

Fig. 5C: Electrophoretic characterization of gp190<sup>S1</sup>  
purified from HeLa cells

Fig. 6A: Expression vector ppT 190 with gp190 sequence

Fig. 6B: Immunofluorescence of the expression of gp190<sup>S</sup>  
in *T. gondii*

Fig. 6C: Polyacrylamide gel electrophoresis of gp190 from  
*T. gondii*

In the gp190/MSP1 precursor protein from *P. falciparum* schematically represented in Fig. 1 the dark blocs stand for regions which are strongly conserved in all strains. The cross-hatched blocs indicate the dimorphic areas, which in the case of the FCB-1 isolate derive from the K1 allele. O1 and O2 indicate the oligomorphous areas. S denotes the peptide signal sequence containing 19 amino-acids, GA the C-terminal region, which includes the signal for the GPI attachment of the protein to the membrane. The arrows indicate the sites of the processing by



which the proteins p53, p31, p36, p30 and p19 arise. The gp190 gene codes for altogether 1639 amino-acids.

The other figures are more conveniently explained in the context of the following Examples.

## EXAMPLES

### Example 1: total synthesis of one of the DNA sequences coding for gp190/MSP1 (see Fig. 3)

#### A. Strategy of synthesis of the gp190/MSP1 gene (gp190<sup>S</sup>) (see Fig. 3A).

The sequence was divided into fragments corresponding to the main processing products: p83, p31, p36, p30 and p19. In the transition regions cleavage sites for restriction endonucleases (arrows in fig.3) were inserted in such a way that the amino-acid sequence was not altered. All the particular cleavage sites are found only once in the sequence.

The fragments were synthesized to overlap, so that the cleavage sites at the respective ends made attachment by ligation to the neighboring fragment possible. All individual fragments contain in addition at their 5' ends a BamHI division site for insertion into expression vectors. The entire sequence could be cloned via MluI and ClaI. The scheme indicated here leads in addition to a sequence which cannot produce the GPI attachment since the C-terminal lacks 18 amino-acids. Synthesis of a corresponding oligonucleotide as well as of a "primer" extending over the SphI cleavage site, leads after PCR to the GA fragment which could be used by SphI and ClaI, the resulting total sequence being gp190<sup>S</sup>. On removing the sequence coding for the peptide signal, "PCR Primer" is produced, over which the fragment  $\Delta S$  has been synthesized. It is permissible to alter the N-terminal via a BamHI and a HindIII cleavage site in such a way that the protein begins with amino-acid no. 20. The nuclear sequence which encodes gp190/MSP1 without signal sequence and without GPI attachment signal was designated gp190S2. Deletion of the GPI attachment signal alone leads to gp190<sup>S1</sup>.



B. Principle of the PCR-supported total synthesis (see Fig. 3B)

Oligodesoxynucleotides of about 120 nucleotides have been synthesized in an alternating manner from the coding or the non-coding strand in such a way that in each case about 20 bases overlapped with the neighboring fragment. The scheme illustrates for example the synthesis of a fragment about 800bp long from oligonucleotides. At the first stage 2 oligonucleotides were amplified "asymmetrically" in each of 4 reaction vessels. This resulted in 4 populations of DNA about 220bp in length, consisting predominantly of single strands (A, B, C, D). Uniting A to B and C to D with amplification over 5 cycles led to 2 approximately 400bp long double-stranded products. Asymmetrical amplification of these DNA fragments (Stage III) resulted in single-stranded populations which following uniting and amplification (Stage IV) resulted after 10 cycles in the end-product G of about 800bp in length. This synthesis could be carried out without isolation of intermediate products and without renewing buffer or enzyme, and was completed in 3 hours. The end-product was purified electrophoretically, divided up with the appropriate restriction endonucleases, and cloned in *E. coli* in pBluescript (Stratagene), to which polylinker a MluI and a ClaI cleavage site had been added.

C. Total sequence of gp190<sup>s</sup> (see Fig. 3C)

Following fusion of all part sequences (Fig. 3A) in pBluescript, the sequence of the gene was checked by the di-deoxy method. The reading frame of gp190<sup>s</sup> had a length of 4917bp (+ 2 stop-codons) and encoded an amino-acid sequence corresponding to that of the gp190/MSP1 from FCB-1 (1639 amino-acids).

D. N- and C-termini of the gp190<sup>s1</sup> variant (see Fig. 3D)

The N-terminal sequence, beginning with the BamHI cleavage site, indicates the transition at amino-acid 20, from which it can be assumed that after splitting of the signal peptide it defines the N-terminus. At the C-terminus the sequence encoded ended at amino-acid 1621. The stop-codon followed the ClaI cleavage site.



## Example 2: Expression of gp190<sup>S2</sup> in E. coli

### A. Expression vector (see Fig. 4A)

The gp190<sup>S2</sup> sequence was inserted via the BamHI and ClaI cleavage sites into pDS56RBSII, by means of which 6 histidines as well as some amino-acids originating in the vector were fused to the N-terminus. This produces the following N-terminal sequence on the reading-frame: Met Arg Gly Ser (His)<sub>6</sub> Gly Ser. Through the promoter P<sub>N25lacD-1</sub> the transcription comes under lacR/O/IPTG control.

### B. Expression and purification of gp190<sup>S2</sup> (see Fig. 4D)

Carrying over the vector pDS56RBSIIgp190<sup>S2</sup> into E. coli DH5alphaZ1 and induction of synthesis through IPTG resulted after electrophoretic separation of the total protein extract from the culture in a clearly visible band of the anticipated size (arrow). Purification of the material through IMAC and affinity chromatography (antibody column with mAK5.2) led to a homogeneous product of about 190 kD. In the Figure M stands for molecular weight standards; 1 = E. coli before; 2 = after induction with IPTG for 2 hours; 3, 4, 5 = fractions from elution of the mAK column.

## Example. 3: Tetracycline-controlled expression of gp190<sup>S1</sup> in HeLa and CHO cells and isolation of the product (see also Fig. 5 and 6c)

A. The gp190 sequence was inserted via the BamHI/ClaI cleavage sites into the expression vector pBi-5. In this way transcription of the gene came under the control of a bidirectional "tTA-reponsive" promoter and could be regulated through Tc. The bidirectional promoter simultaneously initiated transcription of the indicator gene luciferase. In consequence the regulation of the expression could easily be followed (see also Fig. 5A).



#### B. Immunofluorescence of HeLa cells, which express luciferase and gp190<sup>S1</sup> under Tc control

The production of luciferase (left), gp190<sup>S1</sup> (middle) in the absence of Tc was demonstrated in HtTA93-9 cells, which contain the bidirectional transcription unit of (A). Following addition of Tc no noteworthy synthesis of gp190<sup>S1</sup> was shown (as represented in Fig. 5B, right).

#### C. Electrophoretic characterization of gp190<sup>S1</sup> purified from HeLa cells

The HeLa cell clone HtTA93-9 as well as the CHO cell clone CHO27-29 have been cultivated with or without Tc. Cell extracts separated by electrophoresis have been analyzed with mAK5.2 by means of "Western blotting" (Fig. 5C); analysis of the CHO cell line is shown on the left, of HeLa on the right. (1) = culture without, (2) = culture with Tc, (3) = non-transfected HtTA-1 cell line. Molecular weight standards are in each case indicated on the left.

#### D. Purification of gp190<sup>S1</sup> synthesised by HeLa cell clone HtTA93-9

Preparative cultivation of the HtTA line and induction of expression of gp190<sup>S1</sup> by withholding Tc permitted isolation of the gene product by affinity chromatography (mAK5.2 column).

The polyacrylamide gel stained with Coomassie (Fig. 6C) following electrophoresis displayed a product consisting of gp190<sup>S1</sup> as well as another protein of about 50 kD. The latter was not derived from gp190<sup>S1</sup> and thus originated from the HeLa cells. Its projected removal should nevertheless present no difficulty in principle.

#### Example 4: Expression of gp190<sup>S1</sup> in Toxoplasma gondii and purification of the product (see also Fig. 6).

A. The gp190<sup>S</sup> sequence was inserted into the vector ppT via MluI/PstI. This brought the gene under the control of the tubulin promoter ( $P_{\text{tub-1}}$ ) of *T. gondii*. The 3' untranslated region (VTR) originated from the main surface protein of *T. gondii* (SAG-1).



### B. Expression of gp190<sup>S</sup> in T. gondii

Transfection of T. gondii with pTT190 led to the isolation of parasite lines which expressed constitutively gp190<sup>S</sup>. Immunofluorescence with mAK5.2 (middle picture) showed not only expression of the gene but also situated the binding of the expression product close to the surface of the parasite, since it, like SAG-1, provokes the same pattern of immunofluorescence (right section of fig. 6B). On the left in Fig. 6B a phase contrast photograph of the middle picture is shown.

### C. Isolation of gp190<sup>S</sup> from T. gondii.

By means of affinity chromatography (mAK5.2 column) gp190<sup>S</sup> was purified from a prepared quantity of T. gondii ( $5 \times 10^9$  parasites). The extremely pure protein possessed the anticipated molecular weight, as the Coomassie-stained polyacrylamide gel indicated following electrophoresis (2-3 on Figure 6C). At no. (1) on Fig. 6C purified gp190<sup>S1</sup> from CHO cells is represented with molecular weight marked on the left side.

### Example 5: Characterization of gp190<sup>S</sup> with monoclonal antibodies.

The interaction of 16 monoclonal antibodies with gp190<sup>S</sup> from the various heterologous expression systems was reviewed by immunofluorescence on P. falciparum and T. gondii or by "Western blot" on the purified proteins. Complete agreement was found when the two parasites were compared (number of +s indicates the relative intensity of the fluorescence). On Western blotting 12mAK's reacted with gp190<sup>S</sup> from E. coli and T. gondii. On the other hand 3 antibodies did not bind to material isolated from CHO cells. Antibodies 15 and 16, which recognize epitopes from the oligomorphs or the alternative allele (MAD20), did not react. The results are summarized in Table 1, in which ND means "not carried out".



## Example. 6: expression of gp190<sup>S</sup> in heterologous systems

### 1. Expression in E. coli

The gp190<sup>S2</sup> was inserted into the expression vector pDS56, RBSII, where it came under control of the promoter PN<sub>25lacO-1</sub>, which can be controlled via the lac operator/repressor/IPTG system (Fig. 4A). Transfer of the plasmid into repressor-producing E. coli cells, eg E. coli DH5alphaZ1, permitted expression of gp190<sup>S2</sup> under IPTG control. By means of a nickel-chelate column the raw product could be isolated via the N-terminal (His)<sub>6</sub> sequence introduced by the vector. An ensuing affinity chromatography on an antibody column led to an extremely pure preparation. Since the monoclonal antibodies used (mAK5.2) recognized a conformational epitope in the C-terminal region, this 2-step purification selected a full-length intact protein with correct folding at least at the C-terminus (Fig. 4B).

In contradistinction to the natural material the end-product possesses 11 additional amino-acids at the N-terminus, of which 6 are histidines. It contains no N-terminal signal and also no C-terminal attachment sequence. The P. falciparum-specific sequence begins with amino-acid 20 and ends with amino-acid 1621.

### 2. Controlled expression of gp190<sup>S1</sup> in HeLa and CHO cell cultures

The gp190S1 was inserted into the vector pBi-5 and thereby placed under control of a promoter regulable by tetracycline (Tc). The Tc-controlled system was chosen for 2 reasons:

- It belongs to the expression systems with which the highest yield is obtained in mammalian cells.
- Unsecreted foreign proteins at high concentration can interfere negatively with cell metabolism. Synthesis of the desired product is consequently begun only after maturation of the culture.



In the construct pBi5-gp190<sup>S1</sup> a bidirectional promoter was activated by the Tc-controlled transcription activator and initiated transcription of both gp190<sup>S1</sup> and the luciferase indicator gene. In the presence of Tc the promoter is inactive. The transcription unit was transferred into both HeLa and CHO cells, which both synthesize constitutively tTA (HtTA line: Gossen, M. and Bujard, H. (1992), Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA 89, 5547-5551; CHO-tTA line, unpublished). Through cotransfection (Ca<sup>2+</sup>-phosphate method) with a hygromycin-resistance-inducing marker gene was selected for successful chromosomal integration. Hygromycin-resistant clones were then investigated for regulability of the expression >Tc, in which luciferase activity was used as indicator. The gp190 synthesis was tested in well regulable clones (regulation factor •Tc 1000). Immunofluorescence analysis (Fig. 5B) as well as investigation by "Western blot" (Fig. 5C) allowed the identification in both cell types of clones which synthesized gp190 under strictly regulable conditions. The best regulable of 20 clones were in each case subcloned. The subclones HtTA93-9 and CHO27-29 were used for cultures on a scale of 10:1. From cell extracts of these cultures intact gp190<sup>S1</sup> could be isolated by means of affinity chromatography (mAK5.2). The material was homogeneous except for a single cellular component which did not derive from gp190<sup>S1</sup> and made up 25% of the preparation (Fig. 6C). It had to be removed in a further purification step.

### 3. Expression of gp190S in Toxoplasma gondii.

Like *P. falciparum*, *Toxoplasma gondii* belongs to the Apicomplexa and consequently has a protein modification system apparently similar to that of *P. falciparum*. *T. gondii* can be transfected with foreign DNA which is efficiently integrated into the genome and furthermore allows problem-free multiplication of *T. gondii* in cell culture. To obtain a product most like native gp190, gp190S2 is expressed in such a way that the protein is secreted on the surface of the parasite and, as in *P. falciparum*, bound to the membrane via a GPI analogue. In that way the gp190S2 (Fig. 3A) has been inserted (Fig. 6A) into the plasmid ppTMCS (D. Soldati, unpublished) and thereby placed under the control of the *T. gondii* tubulin promoter,

This expression construct was transfected into *T. gondii*. Selection with chloramphenicol led to resistant clones synthesizing gp190 which were detected by immunofluorescence (Fig. 6B).





The immunofluorescence with anti-gp190 antibodies was indistinguishable from a corresponding pigmentation of the parasites by means of antibodies against SAG1, the main surface protein of *T. gondii*. It may be deduced from this that gp190 is bound to the surface of *T. gondii*. Several *T. gondii* clones (Nos. 3.1 to 3.4) were characterized and saved for the production of gp190. Using affinity chromatography (mAK5.2) gp190 was isolated from *T. gondii* cultures (clone 3.4) cultivated on a preparative scale. Electrophoretic analysis revealed a homogeneous product with a migration rate similar to that of the intact protein (Fig. 6C)

Example 7: Characterization of gp190 protein from various expression systems by means of monoclonal antibodies.

A set of gp190-specific monoclonal antibodies, of which a number recognize conformational epitopes, were used to compare the reactivity of the antibodies with *P. falciparum* and *T. gondii* parasites via immunofluorescence. Table 1 shows that the reactivity of the 16 antibodies with either parasite is the same. This is a strong indication that in *T. gondii* "native" gp190 is being mostly produced. Comparison of the reactivity of the antibodies with protein from *E. coli*, HeLa or CHO cells as well as *T. gondii* shows also that most of the antibodies react with the 4 preparations. In particular the protein derived from *E. coli* recognizes more of the antibodies than that produced in mammalian cells. This is apparently a consequence of glycosylation in mammalian cells.

Example 8: Immunization of *Aotus lemurinus griseimembra* monkeys with gp190/MSP from *P. falciparum* (FCB-1).

Two independent immunization experiments (A, B) were carried out. In them in one instance (A) 1.0 mg and in the other (B) 0.6 mg of very pure gp190/MSP1 was extracted from about  $2 \times 10^{11}$  parasites respectively.

The protein was administered together with Freund's Adjuvant (FCA). The control group received only FCA. Immunization equally with the protein mixture or the adjuvant was done three times at intervals of 4 weeks. Two weeks after the last immunization each of the animals



was infected with  $10^5$  parasites (FVO strain) from a donor animal. Parasitaemia was measured daily. The results are summarized in Fig. 2. The symbols mean:

T: that the animals were treated with resochin

D: a dead animal

Fig. 2A: individuals in the vaccinated group each received  
3 x 60 micrograms gp190/MSP1

Fig. 2B: individuals in the vaccinated group each received  
3 x 40 micrograms gp190/MSP1

While in the control group only 1/11 animals did not develop parasitaemia, this was 6/10 in the vaccinated group. The four animals in the vaccinated group who did develop a pronounced parasitaemia did so - in comparison to the control group - with an average delay of four days (exceeding the 2% limit of parasitaemia).

These experiments indicate for the first time a highly significant protection by gp190/MSP1 against infection with *P. falciparum* in a monkey model. The process according to the invention consequently permits a practical vaccine against malaria to be presented for the first time.





Table 1: Interaction of gp 190<sup>s</sup> with monoclonal antibodies

Code	mAb	Type of epitope	Variability	P.f. FCB	IFA		Western blot	
					Toxoplasma	E. coli	Toxoplasma	CHO
1	5.2	conformational	conserved	++++	++++	+	+	+
2	12.10	conformational	conserved	++++	++++	+	+	+
3	7.5	conformational	conserved	++++	++++	+	+	+
4	12.8	conformational	conserved	++	++	+	+	+
5	7.3	conformational	dimorph (K1)	++++	+++	+	+	+
6	2.2	conformational	conserved	++++	++++	+	+	+
7	7.6	conformational	dimorph (K1)	++++	++++	+	+	+
8	9.8	conformational	conserved	++++	++	+	-	-
9	13.2	sequential	conserved	++++	++++	+	+	+
10	13.1	sequential	dimorph (K1)	++++	+++	+	+	-
11	6.1	sequential	dimorph (K1)	++++	++++	+	+	ND
12	A5Z	unknown	unknown	+++	+++	+	+	+
13	17.2	unknown	unknown	++++	+++	ND	ND	ND
14	15.2	unknown	unknown	++++	+++	ND	ND	ND
15	9.7	conformational	dimorph (MAD20)	-	-	-	-	-
16	12.1	sequential	oligomorph	-	-	-	-	-

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A process for preparing the complete gp190/MSP1 protein of *Plasmodium* comprising the step of expressing a gene for gp190/MSP1 protein, wherein in the DNA sequence of the gene has a reduced AT content in comparison to the naturally occurring sequence.
2. A process according to claim 1, wherein the *Plasmodium* is *Plasmodium falciparum*.
3. A process according to claim 1 or claim 2, wherein the expression takes place within a host organism.
4. A process according to any one of claims 1 to 3, wherein the DNA sequence is derived from the DNA sequence of *Plasmodium falciparum* strain FCB-1.
5. A process according to any one of claims 1 to 4, wherein the AT content is reduced from 74% to 55%.
6. A process according to any one of claims 1 to 5, wherein the DNA sequence encodes the complete amino acid sequence including signal peptide and attachment signal.
7. A process according to any one of claims 1 to 5, wherein the DNA sequence encodes the complete amino acid sequence except for the attachment signal.
8. A process according to any one of claims 1 to 5, wherein the DNA sequence encodes the complete amino acid sequence except for the attachment signal and the signal peptide.
9. A process according to any one of claims 1 to 8, wherein the process comprises the following steps:
  - (a) Design of a DNA sequence based upon the DNA



sequence from *Plasmodium falciparum* FCB-1, wherein the DNA sequence takes in to consideration the degeneracy of the genetic code but still maintains the amino acid sequence;

(b) Division of sequence into overlapping regions  
5 wherein regions consist of p83, p31q1, p36, gp30 and gp19;

(c) Synthesis of desoxyoligonucleotides, wherein each of these extend the whole length of a region;

(d) synthesis of the regions coding for gp19, p30, p36, p31 by PCR and synthesis of the region coding for  
10 p83 by fusion of two sequences comprising approximately 1200bp;

(e) individual cloning of coding sequences;

(f) fusion of the complete gene; and

(g) expression in a suitable expression system.

15

10. A recombinant process according to claim 9, wherein the desoxyoligonucleotides synthesized in step (c) should be on average 120 nucleotides long and in each instance overlap the neighbouring sequences by about 20  
20 bases.

11. A recombinant process according to any one of claims 1 to 8, wherein the expression vector is dPS56,RBSII.  
25

12. A recombinant process according to any one of claims 1 to 8, wherein the expression vector is pBi-5.

13. A recombinant process according to any one of  
30 claims 1 to 8, wherein the expression vector is ppTMCS.

14. A recombinant process according to any one of claims 1 to 13, wherein the DNA sequence is expressed in *E. coli*.

35

15. A recombinant process according to claim 14, wherein the *E. coli* strain used is the repressor-producing



strain *E. coli* DH5alphaZ1.

16. A recombinant process according to any one of  
claims 1 to 13, wherein the DNA sequence is expressed in  
5 HeLa cells.

17. A recombinant process according to any one of  
claims 1 to 13, wherein the DNA sequence is expressed in  
CHO cells.

10 18. A recombinant process according to any one of  
claims 1 to 13, wherein the DNA sequence is expressed in  
*Toxoplasma gondii* or *Leishmania*.

15 19. A complete DNA sequence which encodes the  
gpl90/MSP1 surface protein of *Plasmodium*, wherein the DNA  
has a reduced AT content compared to the naturally-  
occurring DNA sequence.

20 20. A complete DNA sequence according to claim 19,  
wherein the *Plasmodium* is *Plasmodium falciparum*.

21. A DNA sequence according to claim 19 or claim 20,  
wherein the DNA sequence is derived from the DNA sequence  
25 of *Plasmodium falciparum* strain FCB-1.

22. A DNA sequence according to any one of claims 19  
to 21, wherein the AT content is reduced from 74% to 55%.

30 23. A DNA sequence according to any one of claims 19  
to 22, wherein the DNA does not code for the attachment  
signal.

24. A DNA sequence according to claim 23, wherein the  
35 DNA does not code for the signal peptide.

25. A DNA sequence according to claim 24, wherein the



DNA includes a sequence which encodes 11 additional amino-acids, of which 6 are histidines, present at the N-terminus.

5 26. A DNA sequence according to any one of claims 19 to 25, wherein the sequence includes no recognizable "splice donor" and "splice acceptor" signals.

10 27. A DNA sequence according to any one of claims 19 to 26, wherein the sequence includes no large GC-rich sequences.

15 28. A DNA sequence according to any one of claims 19 to 27, wherein the sequence includes no recognition signals for restriction enzymes recognizing sequences of six or more base pairs.

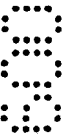
20 29. A DNA sequence according to any one of claims 19 to 28, wherein the sequence for recognition signals of particular restriction nucleases in regions which, after processing of the protein, separate existing domains, contains uniquely occurring cleavage sites for restriction endonucleases.

25 30. A DNA sequence according to any one of claims 19 to 29, wherein the sequence has at both its ends cleavage sites for endonucleases which do not appear in the remaining sequence and in a vector for utilization.

30 31. A host organism which contains the complete DNA sequence according to any one of claims 19 to 30.

35 32. A host organism according to claim 31, wherein the organism is *E. coli*.

33. A host organism according to claim 32, wherein the *E. coli* strain is the repressor-producing *E. coli*



strain DH5alpha21.

34. A host organism according to claim 31, wherein the host organism is HeLa cells.

5 35. A host organism according to Claim 31, wherein the host organism is CHO cells.

10 36. A host organism according to claim 31 or claim 32, wherein the host cells synthesize constitutively tTA.

15 37. A host organism according to claim 31, wherein the host organism is selected from the group consisting of *Toxoplasma gondii*, *Leishmania*, *baculovirus*, *adenovirus* and yeast.

20 38. Use of a gp190/MSP1 protein produced by a process according to any one of claims 1 to 18 for active immunization against malaria.

39. Use of a gp190/MSP1 protein produced by a process according to any one of claims 1 to 18 for producing monoclonal antibodies suitable for passive immunization.

25 40. Use of a DNA sequence produced by a process according to any one of claims 19 to 30 for producing a vaccine based on nucleic acids.

30 41. A process for the stabilization of gp190/MSP1 gene sequence from *Plasmodium*, comprising the step of reducing the AT content of the sequence compared to the naturally-occurring sequence.

35 42. A stabilized *Plasmodium* gp190/MSP1 gene having a reduced AT content compared to that of an unstabilized gene.





43. A vector containing a DNA sequence according to any one of claims 19 to 30 and/or claim 42.

44. A host cell containing a vector according to claim 43.

45. A vaccine containing a protein produced by a process according to any one of claims 1 to 18 and/or a DNA sequence according to any one of claims 19 to 30 and/or a host according to any one of claims 31 to 37 and/or a vector according to claim 43.

46. A vaccine according to claim 45, further comprising immunity-promoting products from *Plasmodium*.

47. A vaccine according to claim 46, wherein the *Plasmodium* is *Plasmodium falciparum*.

48. A process according to claim 1 substantially as hereinbefore described with reference to any one of the examples.

Dated this 5th day of October 2000

HERMANN BUJARD

By their Patent Attorneys

GRIFFITH HACK

Fellows Institute of Patent and  
Trade Mark Attorneys of Australia



### Summary

Recombinant manufacturing process for a complete malaria antigen gp190/MSP1

The invention concerns a recombinant manufacturing process for the complete gp190/MSP1 surface protein of Plasmodium, in particular *P. falciparum*, as well as the complete DNA sequence of this protein and suitable host organisms for expression of the sequence, by means of which the protein in its entirety can for the first time be synthesized outside the parasite.

The invention opens up for the first time the possibility of manufacturing the gp190/MSP1 surface protein in sufficient quantity; in addition there is the matter of making gp190/MSP1 available as a vaccine.

Finally the invention sets out a process for the stabilization of AT-rich genes (Fig. 1).



Fig. 1

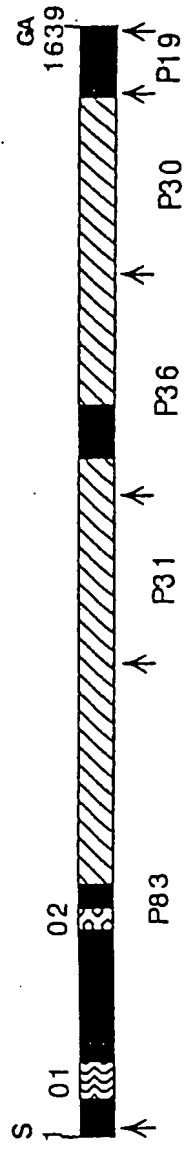


Fig. 2

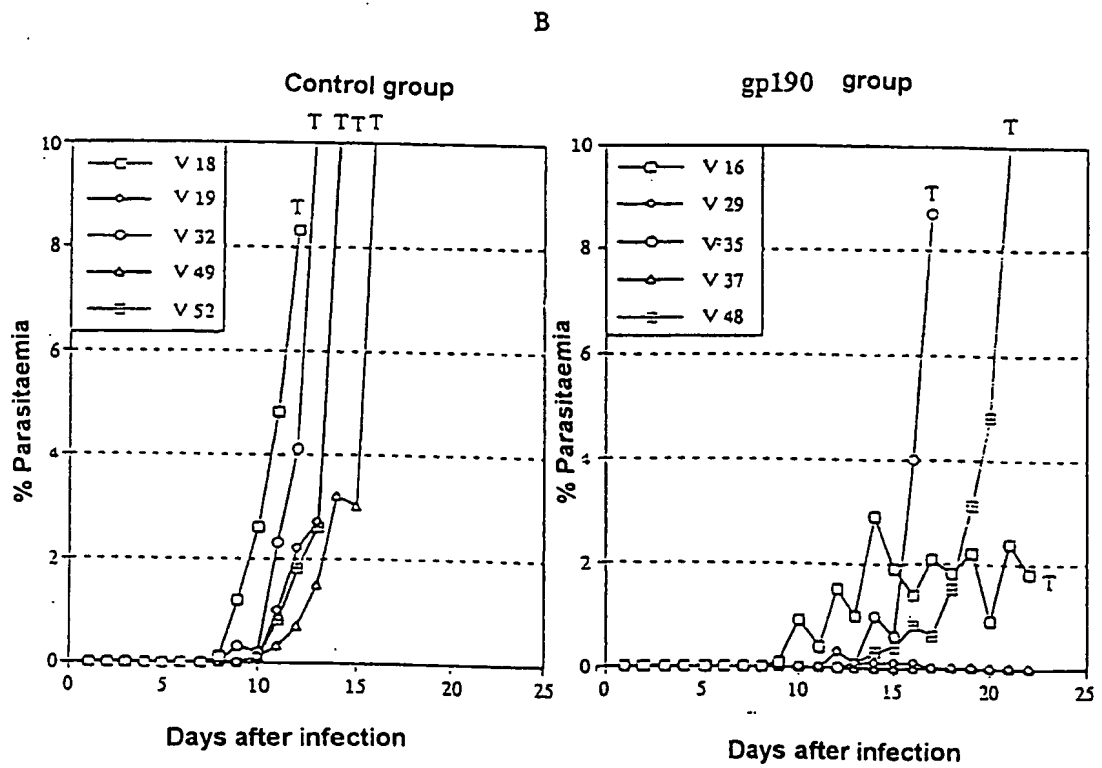
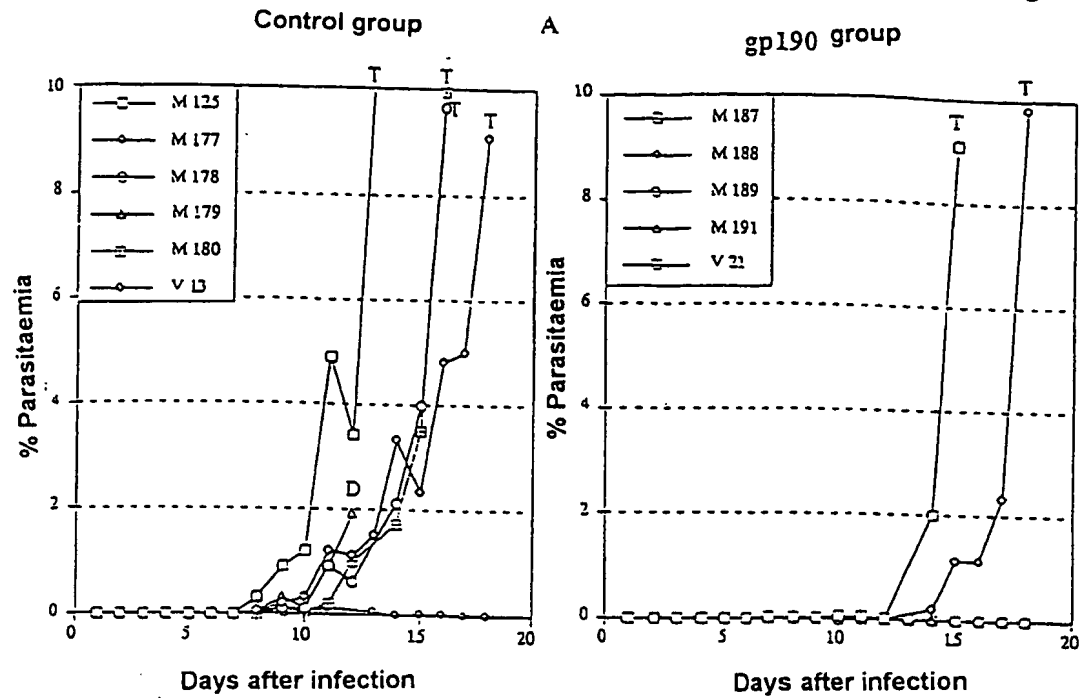


Fig. 3A

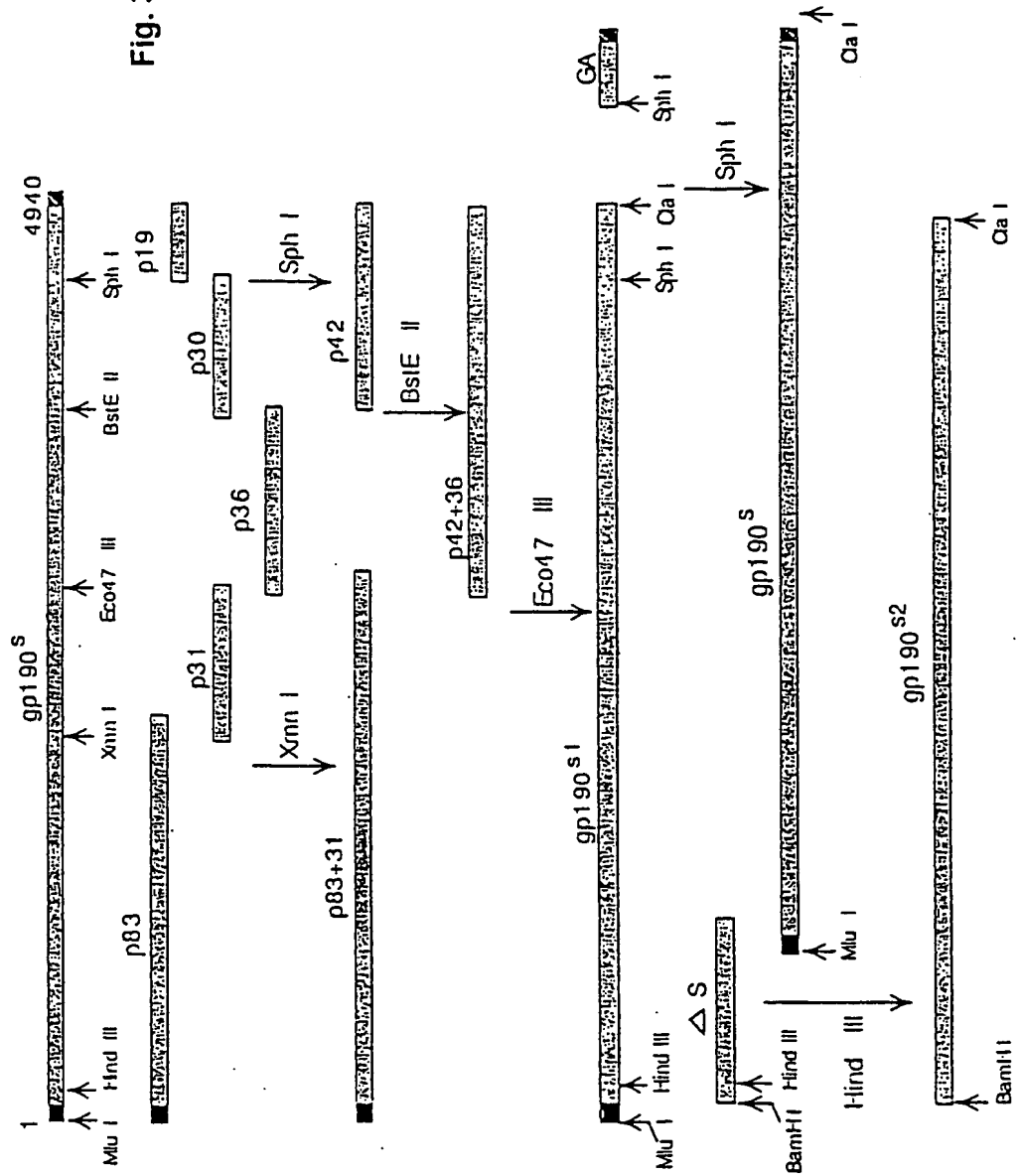
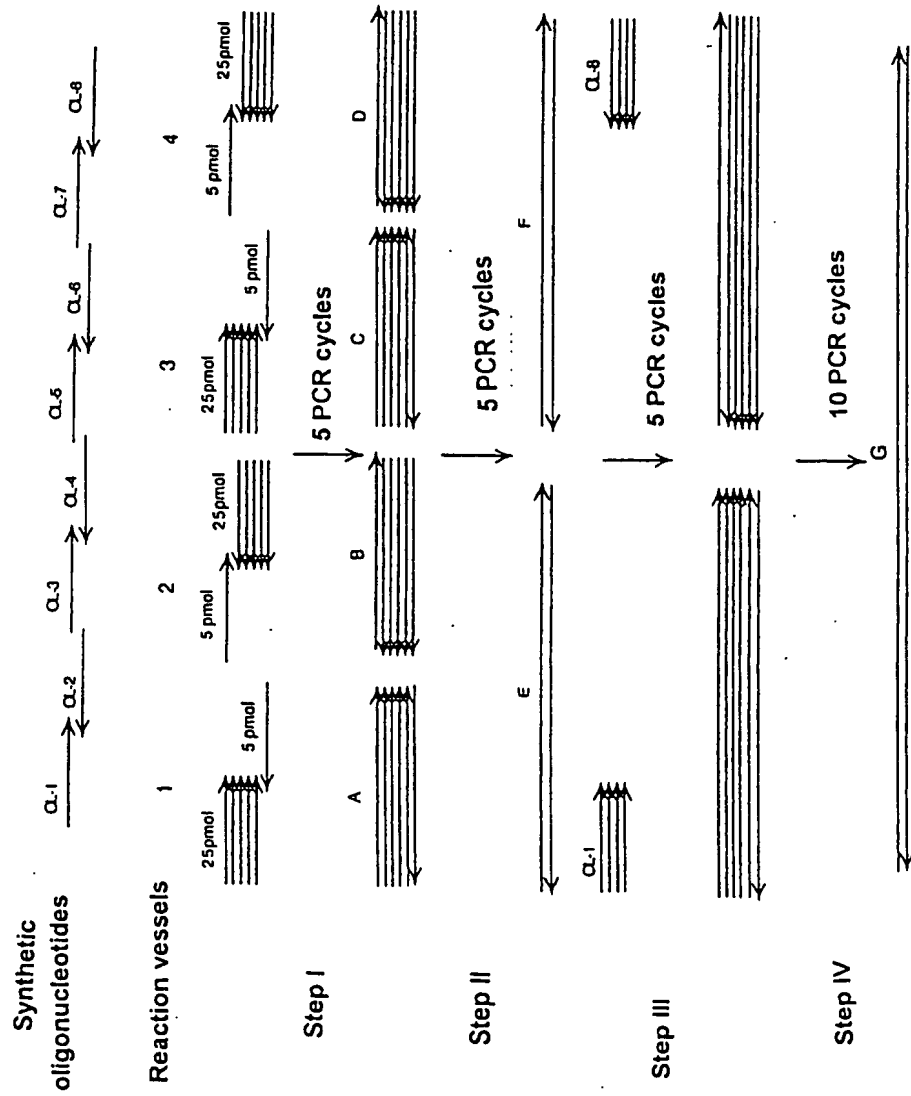


Fig. 3B



DNA sequence of the native (gp190n) and of the synthetic gene (gp190s) for gp190 from FCG-1

```

AS      M K I I F F L C S F L F F I I N T O C V T H E S Y Q E      27
gp190n      G A T T A T T A A A T A A T A A T A A A
gp190s      C C C A C C C G A T G A A A A T C A T T T C T C T G T T C A T T T C T G T T T T A T C A T C A A T A C T C A G T G C C T G A C C C A C G A A T C C A T C A G G A G      90
              R L U 1

AS      L V K K L E A L E D A V L T G Y S L F O K E K M V L N E G T      57
gp190n      T C A A A A A A T G A T T T T A T A A A A T A A A A
gp190s      C T G T T A A G A A A C T G G A A C T T T G A A G A T G C C T T A C C G G A T A C A G C C T T T C C A G A A G G A A G A T G T G C T G A A T G A A G G A C G      180

AS      S G T A V T T S T P G S K G S V A S G G S G G S V A S G G S      87
gp190n      A A T T T A C T A T T C A T A C A T T A T C A
gp190s      A G T G C A C G C C C T T A C A C G A C A C C C G T T C A A A G C G T C T G C C T A G C C T G C C T G C C G T C T G C C C T C T G C C G T T C C      270

AS      V A S G G S V A S G G S V A S G G S G N S R R T N P S D N S      117
gp190n      T T A T T C A T T T T T C A T A T T T C A A C T A T A T T A T T A T T A T T A T T A
gp190s      G T C G C C T C C C G C A C G T G C C A C C T G C C T C A G T G C C A A C C G C C G T T C C C G A A C A G T G A A G A A C A A T C C A T C T G A C A A C T C T      360

AS      S D S D A K S Y A D L K B R V R N Y L L T I K E L K Y P O L      147
gp190n      T A T T A T T T A A A A A C T C T G T A A A C A T T A C C
gp190s      A G C G A T T C C G A C C C A A G T C T A C C G C A C C T C A A G C A C C A G T G A G A A C T A T C T C T C A T A T C A A G A G C T G A G T A C C C A C A G T T G      450

AS      F O L T N B H L T L C D N I E G F K Y L I D G Y E E I N E L      177
gp190n      T T A T T A T T T T T T T T T T T T T T A T T A T T A
gp190s      T T C G A C C T C A C T A A T C A T C T G A C T G T C T A A C A T T C A T G C C T C A A A T A T C T G A T T G A C C G T A C G A A G A T C A A T G A A C T C      540

AS      L Y K L N F Y F O L L R A K L N D V C A N D Y C Q I P F N L      207
gp190n      T A T A A C T T T T A T A A T A T A T T A T T T C T
gp190s      C T G T A C A G T G A A T T T C A C T T C G A C T T C A A G G C C A A A C T G A A T A C G T T T C C C C A A G A C T A T T G T C A A T T C C A T T C A A T T T G      630

AS      K I R A N E L D V L K K L V F G Y R K P L D N I K D N V G K      237
gp190n      A T C T A T A A A C T A A C T G A A A A T A T T A T T A A
gp190s      A A G A T C A G A C C A A G A G T G C A C C T A T T G A A G A A T T G C T T T C G A T A T C C G A A G C C T C T G A C A A C A A T G T G C C A A A G      720

AS      H E D Y I K K N K K T I E N I N I L I E E S K K T I D K N K      267
gp190n      C A A A A A T A T A T A G T G A A T T
gp190s      A T G C A A G A T T A T T A A A A G A A C A G A C A T C G A A C A T T A A C A G C T G A T C G A A G A T C C A A A G A C C A T A G A C A A A A T A A G      810

AS      N A T K E E E K K R L Y Q A O Y O L S I Y N K Q L E E A N      297
gp190n      T A A A A A A A T A T T T T T T T T T C T A T A A T
gp190s      A A T G C A A C A A G G A G A A A A G A A G A G T T G A C A G C C C A G T A C C A C C T C T C A T C T A T A A A A C A G C T T G A A G A G C C A T A A C      900

AS      L I S V L E K R I D T L K X N E N I K E L L D K I N E I K N      327
gp190n      T A A T A A A T T T T A A A C T G T A T T A
gp190s      C T C A T C A G C T A C T G A A A C C C A T A G A C A C C T C A A G A A A T C A A A T A T C A A G A A C T C T C G A C A G A T A A T G A A T T A A G A A T      990

AS      P P P A N S G N T P N T L L D K N K K I E E E E K E I K E I      357
gp190n      C A G T A T A T A A T T C T T A A C A A A A A A A A T
gp190s      C T T C C C C A G C C A A C T T G G A A C A C C C T A A C A C C T G C T G C A A G A C A A G A T A G A G C A C C A G A A A G A G A T C A A A G A G A T C      1080

AS      A K T I K F N I D S L F T D P L E L E Y Y L R E K N K N I D      387
gp190n      T A T T T A G T A A A A A T A A A A A A A A T A G A C
gp190s      G C C A A A C C A T T A A G T T C A C A T A G A T T C T C T T A C T G A T C C C C T G A C T G A G T A C T A C T G A C A G A G A A T A A G A A T A T A G A C      1170

AS      I S A K V E T K E S T E P N E Y P N G V T Y P L S Y N D I N      417
gp190n      A A G T A G T A T C A A A A A A A T T T T A
gp190s      A T C T C C C A A A G T C G A G A C A A A G A A T C A A C G A A C C T A A T G A A T A T C C A A T G G T T G A C G T A C C C T C T G T C T A A C G A T A T C A A C      1260

AS      N A L N E L N S F G D L I N P F D Y T K E F S K N I Y T D N      447
gp190n      T T A T A T T C T T T A T A T A T A A A A C A C A T T T
gp190s      A A C C C T C A A C G A C C T A A T A G C T T C G T G A C T T G A T A A C C C C T T C A T A T A C A A A G A A C C C T C T A A G A A T A T C A C A G A C A A T      1350

AS      E R K K F I N E I K E K I K I E K K K I E S D K K S Y E O R      477
gp190n      A A A C A T T A A A T A A A A A A A A T C T A T C A A
gp190s      G A G A A A G A A C T T A T C A A G A A A T C A A G A A G A T C A A A T T C A G A A G A A A A T T C A G A G T G A C A A G A A A A G T T A C G A A G A C C C      1440

AS      S K S L N D I T K E Y E K L L N E I Y D S K F N N N I D L T      507
gp190n      T C T G T C T A A A A A A T A T T A T A T A G T T A T A T
gp190s      A G C A A A A G T T A A A C G A T A C A T A A A G A T C A A A A G C T G A A C G A G A T C T A T G A T T C A A A T T C A A C A A T A C A T G A C C T G A C C      1530

AS      N F E K H H G K R Y S Y K V E K L T E B N T F A S Y E N S K      537
gp190n      T A T A A C A T T T T T T T T T T A A T A A
gp190s      A A C T T C G A A A A T G A T G C A A A C C T A C T T T A C A A A G T G C A A A C T G A C A C C A T A A T A C C T T G C A T C C T A C A A T T C T A A G      1620

```

\_\_\_\_\_



AS L E K K K L S Y L S S G L B E L I A E L K E V I K N K N Y T 1167  
 gp190<sup>n</sup> T A A A A T A T C A T A A T T A T T A T T A T A A A A A T A T T A  
 gp190<sup>s</sup> C T G G A G A G A G A G C T C A G C T A C T C T A G C C G A C T C A T C A C C T C A T C C C G A G C T C A G C A G T C A T T A A G A C A G A C T A C A C C 3510  
 AS G N S P S E N N T D V N H A L E S Y K K F L P E G T D V A T 1197  
 gp190<sup>n</sup> T T C T A G T T C T T A G T T C T T A A A T C A T A A A T C A T A A  
 gp190<sup>s</sup> G G C A A T A G C C A A G C G A A T A T A C A G A C T G A A T A C G A C T G G A A T T A C A G A A G T T C T C C T G A G G A C A G A T G T C C C A C T 3600  
 AS V V S E S G S D T L E O S O P K K P A S T E V G A E S N T I 1227  
 gp190<sup>n</sup> T A A G A G A T A A A A G A A A A A A A A A A A A T C A  
 gp190<sup>s</sup> C T G G T C T G A A T C G G T C C G A C A C A C T G A G C A G T C A A C T A A G A G C C T G C A S T A C T C A T C T C G G A G C C A G T C C A C A C A A T 3690  
 AS T T S O N V O D E V D D V I I V P I T G E S E E D Y D D L G 1257  
 gp190<sup>n</sup> A A A A T T A A A A A A A A A A A A A T A A T C A A A T T T A A  
 gp190<sup>s</sup> A C C A C A T C T C A G A A C T C G A C G A G A G G T C A T G A C T C A T C A T T G T G C C T A T C T T C G G C A G A G G A G G A G A C T A C A T G A C C T C G C 3780  
 AS Q V V T G E A V T P S V I D N I L S K I E N E Y E V L Y L K 1287  
 gp190<sup>n</sup> A A A A A A A A A A A A A A T T T A T T G T T A T  
 gp190<sup>s</sup> C A G G T C C T C A C C C T C A G C C T C A C T C C T C C G T G A T G A A A C A S T T C T G C A A A T C G A G A C G A A T A C G A A C T G C T C T A T C T A A 3870  
 AS F L A G V Y R S L K K Q L E N N V M T F N V N V E D I L N S 1317  
 gp190<sup>n</sup> T A T T A A G T A A A T A A A T A T T T T T T T A T T C A  
 gp190<sup>s</sup> C C T C T G C A G C C C T C A T A G C T C T C A G A A A C A C T G A G A A A A C C T G A T G A C C T C A A T G C A C T G A A G C A C A T C T C A A C A C G 3960  
 AS R F N K R E N F K N V L E S D L I P Y K D L T S S N Y V V K 1347  
 gp190<sup>n</sup> A A C T A A T T A A T C A T A A A T T A A A A A A A A A A A  
 gp190<sup>s</sup> C G C T T A A T A G A G A G A A A A T C A A G A A C C T T T G A G A C C G A C T T G A T T C C T A T A A G A C C T G A C C T C T T A A C T C T C A A G 4050  
 AS D P Y K F L N K E K R D K F L S S Y N Y I K D S I D T D I N 1277  
 gp190<sup>n</sup> T T A T T A A A A A C T A A G C T T T T A A T G A  
 gp190<sup>s</sup> G A C C A T A C A G T T C C T C A T A A A G A G A G C C A A A T T C T G T C A G T A C A A C T A T A T C A G A C T C C A T G C A C C A T A T C A A T 4140  
 AS F A N D V L G Y Y R I L S E K Y K S D L D S I K Y I N D K 1407  
 gp190<sup>n</sup> T A T T A  
 gp190<sup>s</sup> T T C C T A A T G A T G T C G G G T A T A C A G A T C C T G A G C A A A A T A C A A G T C T G A C T T G A C T T A T A A A A G T A T A T A A C G A T A A G 4230  
 AS G G E N E K Y L P F L N N I E T L Y K T V N D K I D L F V I 1437  
 gp190<sup>n</sup> T A C C T T A C T T G T A T A T T T T T T T T T T T T T T A T  
 gp190<sup>s</sup> C A A G C C A A T G A A A A A T C T G C C C T C C A A T A C A T C G A A A C C T T T A C A A C A G A C T G A C C A A A A T C G A C T C T T C T G A A T 4320  
 AS E L E A K V L N Y T Y E K S N V E V K I K E L N Y L K T I O 1467  
 gp190<sup>n</sup> T T A A A A T A T A T A T A T C A C A A A A A T T T A T  
 gp190<sup>s</sup> C A C T G C A G C C A A G C C T C C A C T A T A C T T A C G A G A G C A A T T G C A G T T A A A A T C A A G A C T C A A C T A C C T A A A A C A A T C C A A 4410  
 AS D K L A D F K K N N N F V G I A D L S T O Y N E N N L L T K 1497  
 gp190<sup>n</sup> A T T A T T T A A A A T T C T A C  
 gp190<sup>s</sup> G A C A A G C T G C A G A T T C A A G A A A A T A A C A A T T C T C G G A A T T G C A G A C T G T C T A C C G A T A T A C C A C A A A T C T C T C A C A A G 4500  
 AS F L S T G H V F E N L A K T V L S N L L D G N L Q G H L N I 1527  
 gp190<sup>n</sup> C T A G T A T T T T T T C T T A T T A T A T A T A T  
 gp190<sup>s</sup> T T T C T G C A C T G C C A T G G T T T C A A A C T G C C A A A C A G T C C T C A C C A A T T C T C G A C C C A C C T G C A G C C A T C T G A A C A T C 4590  
 AS S O E Q C V K K Q C P O N S G C F R E L D E R E E C K C L L 1557  
 gp190<sup>n</sup> A  
 gp190<sup>s</sup> T C C C A G C C A A T G C C G T A A G A A A C A G T C C C C A G A A T A G C G G C T G T T C A G G C A T G T G A C C A G C C A A G A G T G C A A G T G T C T C T G 4680  
 AS N Y K Q E G D K C V E N F N P T C H E N N G G C D A D A K C 1587  
 gp190<sup>n</sup> T A T T A T T T T C T T A T A C T  
 gp190<sup>s</sup> A A C T A C A A C A A G A G A T A A G T C C T G A G A C C C A A C C T A C T C A A T G A A A A C A A T G C C G G T G T G A C C C G A T C C A A T G C 4770  
 AS T S E D S G S N G K R I T C E C T E P D S Y P L F D G I F C 1617  
 gp190<sup>n</sup> A T T C A T T T C  
 gp190<sup>s</sup> A C C G A G A A C A C C C C C T A A C G A A A G A A A T C A C A T G C G A G T G T A C T A A G C C G A C T C C T A T C A C T C T T G A C C G G A T T T T T C C 4860  
 AS S S S N F L G I F Y L L I L M L I L Y S F I \* 1639  
 gp190<sup>n</sup> A C T T C C T A A A C A T A T A A T A A T A T  
 gp190<sup>s</sup> T C A G C C T A A T T C C T G C C A T T C T T C C T G A T T C A T G C T G A T C T G T A C A C C T T C A T T A A T A G A T T G A T C 4940  
 stop codon C1a 1

Fig. 3D

	N'-terminus	C'-terminus
gp190s1 Sequence		
DNA Sequence	GC <u>ACGCGTATG</u> AAAATC	AGCTCTAATTAATAGGCGGCCGCATCGATGGC
AA Sequence	Mlu-I Met Lys Ile	Ser Ser Asn stop codon
AA Position	1 2 3	1619 1620 1621
gp190s2 Sequence		
DNA Sequence	GC <u>GGATCC</u> GTGACCCAC	AGCTCTAATTAATAGGCGGCCGCATCGATGGC
AA Sequence	BamHI Val Thr His	Ser Ser Asn stop codon
AA Position	20 21 22	1619 1620 1621

Fig. 4A

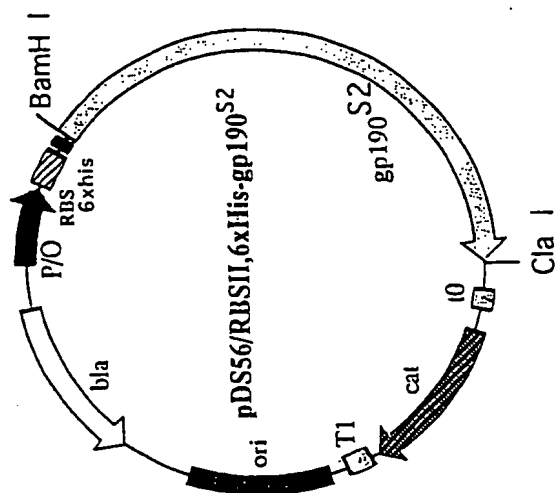


Fig. 4B

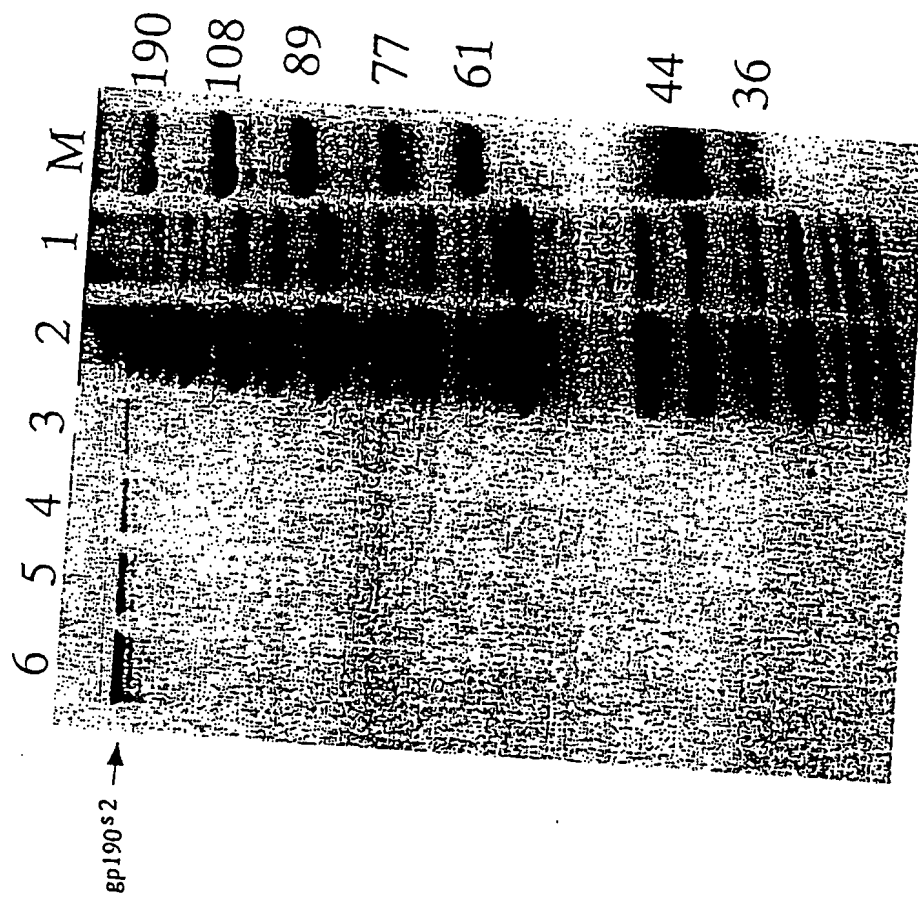


Fig. 5A

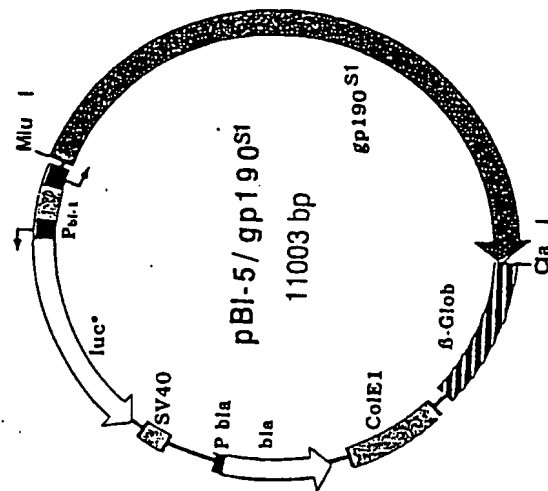


Fig. 5 B

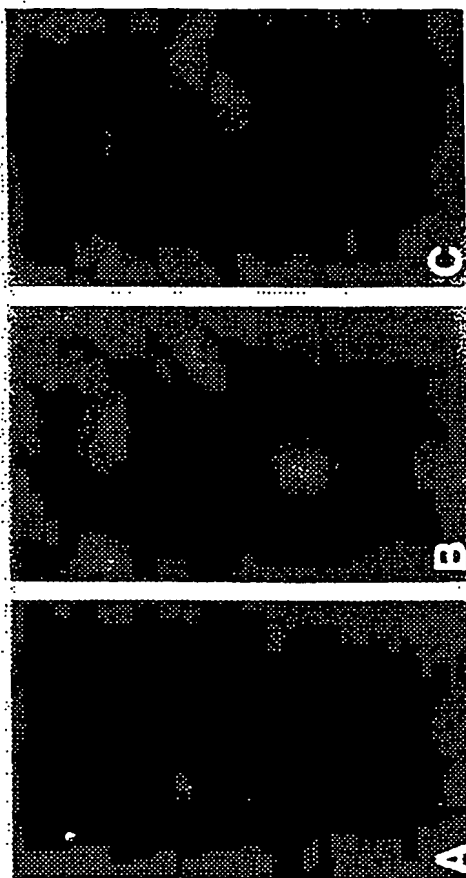


Fig. 5C

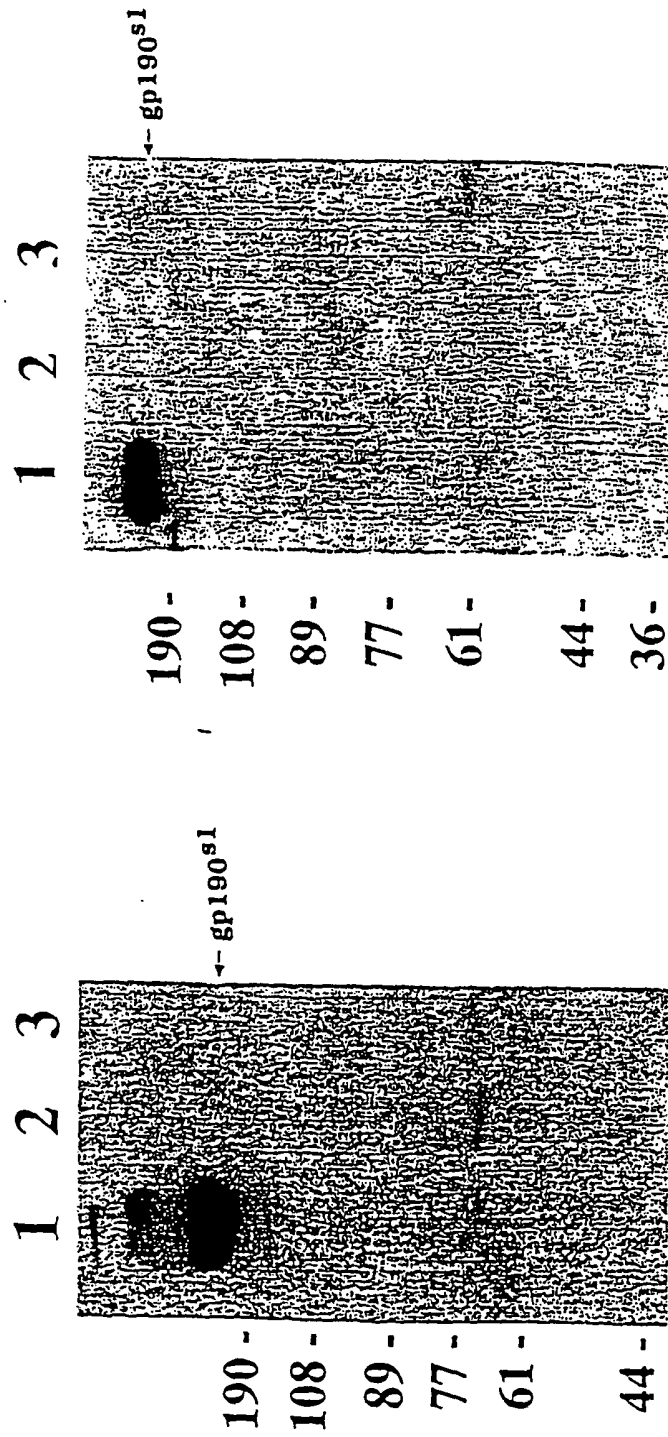


Fig. 6A

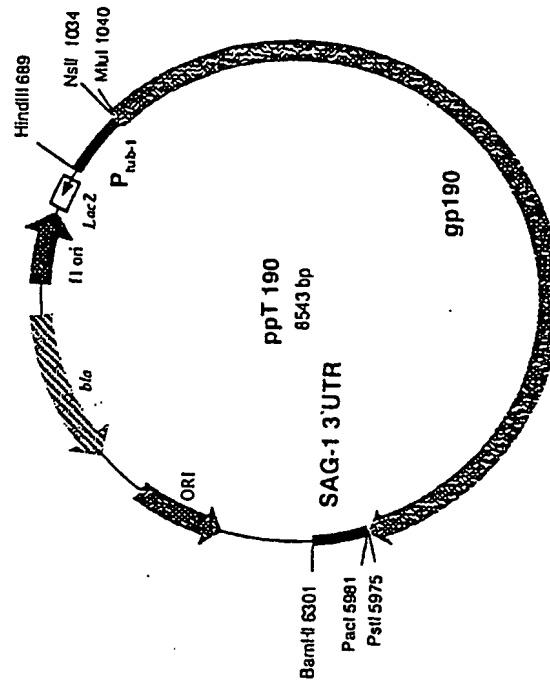




Fig. 6B

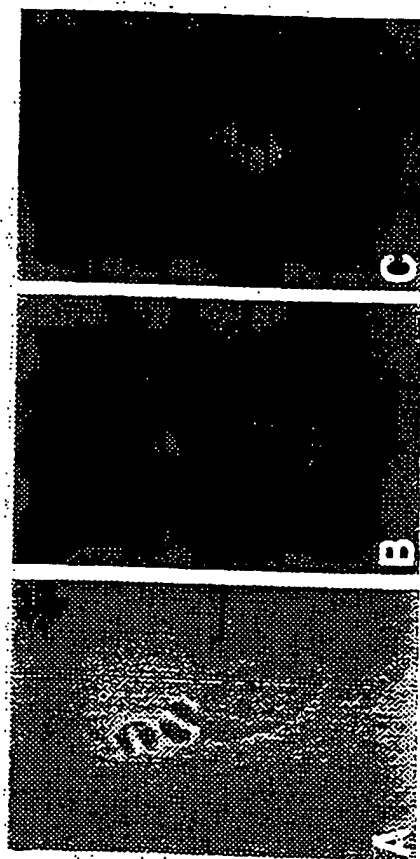
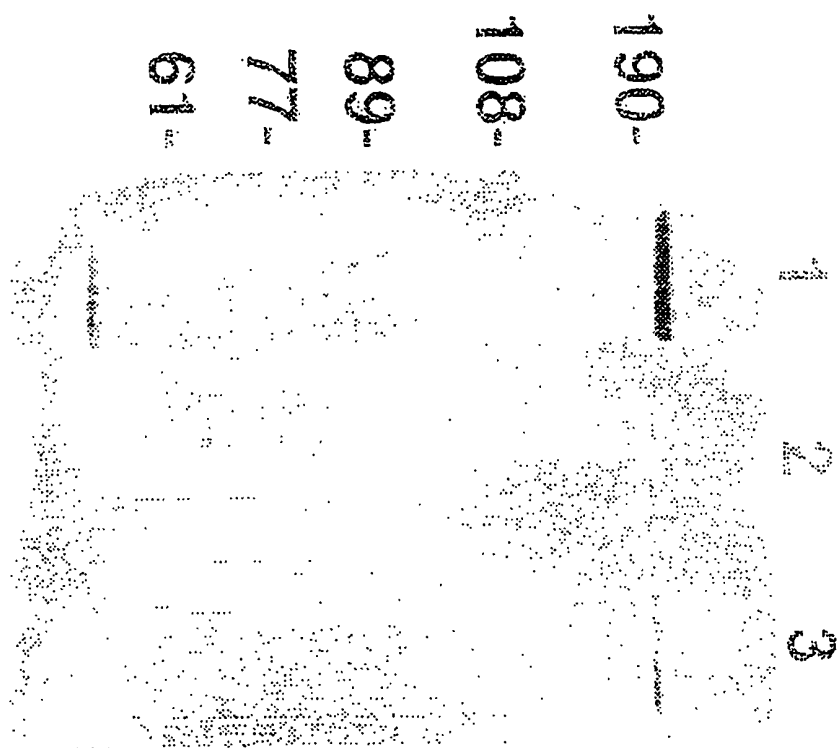


Fig. 6c



**(12) PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

**(11) Application No. AU 199748649 B2**  
**(10) Patent No. 727864**

**(54) Title**  
**Meth d for producing recombinants intended for use in a complete malaria**  
**antigene GP190/MSP1**

**(51)<sup>7</sup> International Patent Classification(s)**  
**C12N 015/30 C12N 001/11**  
**A61K 031/70 C12N 001/21**  
**A61K 039/015 C12N 005/10**  
**C07H 021/00 C12N 015/62**  
**C07K 014/445 C12N 015/67**

**(21) Application No: 199748649**

**(22) Application Date: 1997.10.02**

**(87) WIPO No: WO98/14583**

**(30) Priority Data**

<b>(31) Number</b>	<b>(32) Date</b>	<b>(33) Country</b>
<b>19640817</b>	<b>1996.10.02</b>	<b>DE</b>

**(43) Publication Date : 1998.04.24**

**(43) Publication Journal Date : 1998.06.18**

**(44) Accepted Journal Date : 2001.01.04**

**(71) Applicant(s)**  
**Hermann Bujard**

**(72) Inventor(s)**  
**Hermann Bujard; Ralf Tolle; Weiqing Pan**

**(74) Agent/Attorney**  
**GRIFFITH HACK,GPO Box 1285K,MELBOURNE VIC 3001**